Spring 5-15-2022

Cellular Automata: The Mathematical Theory Behind the Spread of COVID-19 and Prediction of Future Spread

Alison Drumheller
University of Lynchburg, drumheller_ar@lynchburg.edu

Recommended Citation

Cellular Automata: The Mathematical Theory Behind the Spread of COVID-19 and Prediction of Future Spread

Alison Drumheller

Senior Honors Project

**Submitted in partial fulfillment of the graduation requirements
of the Westover Honors College**

Westover Honors College

April, 2022

_____
Leslie Hatfield, PhD

_____
Kevin Peterson, PhD

_____
Christine H. Terry, PhD

Project Summary

This thesis seeks to analyze the spread of the original COVID-19 strain and develop a mathematical model to predict the chances of being infected by this disease using a number of variables. This model is based on the mathematical theory of cellular automata, otherwise known as the theory of spread. The research uses real world data of COVID-19 which includes infection rate, death rate, vaccination rate, use of masks, and transmission rates. By using cellular automata, we predict that the more preventative measures an individual puts in place for themselves, the less likely they are to be infected by the virus. Therefore, COVID-19 data will be used to calculate the likelihood of infection based on certain factors, such as if the person is masked and/or vaccinated, in a set environment, such as a classroom with a specified number of people around. The results from this research will predict the best way for a person to limit their chances of contracting the virus. It will also give us results on the optimal combination of factors that will be most effective at limiting the spread.

Keywords: COVID-19, cellular automata, infection rate, vaccination, masks, transmission rate, spread of disease

Introduction

In early 2020, the COVID-19 virus ran rampant throughout the entire world, causing a global pandemic. Many people got infected and, in severe cases, died from this virus. Countries across the world attempted to eradicate the virus by asking citizens to wear masks and get vaccinated once vaccines were made available. Governments even put entire cities and states through a lockdown with the hope of limiting people's exposure to this deadly disease. It was incredibly important to come up with some sort of answer as to how to combat COVID-19. Studies have shown that when it comes to deadly, easily transmitted diseases (such as HIV), it only takes a handful of infected people to cause a major outbreak (Ramos & Schimit, 2019). Many studies were conducted to determine which preventative measures were most effective. This included testing different theories and formulas, which were used to develop a model that tracks the spread of disease. A mathematical model can be used to determine how quickly COVID-19 spreads from person to person, as well as the likelihood of becoming infected once someone comes into contact with another person who is infected. One of the most popular theories involved with the theory of disease spread, especially in the realm of mathematics, is cellular automata.

Cellular automata, also known as the theory of spread, was made popular by Stephen Wolfram's analysis in his book *A New Kind of Science.* Cellular automata is a model that looks at individual cells. Different rules are set within the model, which impact how the individual cells react towards each other and change (Wolfram, 2002). This essentially means that a cellular automata model closely investigates the relationship between neighboring cells and interprets the

results. These models are often used in order to determine the spread of disease, because in order for someone to get a disease, they must come in contact with someone who is infected.

Many different models of cellular automata have been used to predict and/or accurately model the spread of an epidemic. For example, a model was created to determine how the disease spread from person to person during a general epidemic, such as the Spanish Flu or the Severe Acute Respiratory Syndrome (SARS) pandemics (White et al., 2007). Certain parameters, included in this study, are important to acknowledge, such as gaining immunity once becoming infected with the disease. Also, the population within the model must be allowed to move from cell to cell in order to simulate human traffic patterns. While this model is based on a very small portion of the population, a 50 by 50 cell array, it is an accurate representation of how disease can spread from person to person. Studies like this are the basis for this project, which applies cellular automata to COVID-19.

In order to model the spread of COVID-19, it is important to understand how the disease spreads from person to person. Tanu Singhal, a well-known pediatrician, describes COVID-19 as being a highly contagious disease transmitted through infected droplets which, if not treated quickly enough, can transition into pneumonia or, even worse, multi organ dysfunction (Singhal, 2020). This disease can be very debilitating for those who get it and, in some cases, can be life threatening. These are just some of the reasons as to why different measures, such as wearing masks, were put into place to curb the spread of infection as much as possible. These preventative measures were especially important when it came to reopening schools after the shutdown at the beginning of the pandemic.

There have been many different studies done in order to determine the best ways to reopen schools. One of these studies looked at how reopening college campuses impacts both the

campus and the surrounding community (Benneyan et al., 2021). They analyzed various reopening measures to determine what the safest method is from a health perspective. The results showed that universities with some form of reopening precautions had approximately a 25% decrease in infection amongst their student body, in comparison to schools who did not have any precautions (Benneyan et al., 2021). This shows that implementing mask and vaccination policies do have a positive influence on the overall rate of infection that occurs not only on a college campus, but throughout the entire population.

COVID-19 can spread throughout a classroom setting extremely quickly. An individual-based model examined how clusters that form in schools/classrooms and control measures can impact the spread of COVID-19. The results showed that large groupings of students correlate to a large spread of COVID-19. This shows that schools should have additional preventative measures for their student body other than testing those for COVID-19 after they begin to show symptoms (Tupper & Colijin, 2021). Different preventative measures were suggested based on the individual-based model, including reducing community transmission, regular and frequent testing of all students (not just those showing symptoms) and taking steps to prevent clusters (or crowds) from forming (Tupper & Colijin, 2021). All of these measures are important in regards to preventing COVID-19 from spreading throughout an entire school's population.

However, it is important to note that no matter how many preventative measures are put into place, such as being vaccinated and wearing masks in indoor settings, people may not always comply. In our model, we look at how the combination of being masked/unmasked and vaccinated/unvaccinated impacts everyone in a classroom setting. We will also use these parameters to predict who is most likely to become infected and the extent of an outbreak over a

specified period of time. In order to understand how well our model works, we will compare our results to studies that use similar methods to determine the probability of a person becoming infected depending on who they have been exposed to.

Ghosh and Bhattacharya (2021) used a spatially explicit epidemiological model composed of five different subpopulations (susceptible, recovered/removed, exposed and asymptomatic, shown symptoms but not positive, and quarantined/hospitalized) to predict the spread of COVID-19. Their results showed that, for both the initial infection level and the probability of becoming infected, there was a bell-shaped curve that occurred. This means that, at first, the risk of infection increased dramatically. However, once infection hits a peak, the risk begins to decrease until there is little to no threat of coming in contact with the virus. These results are very similar to another study using cellular automata to map the spread of the ebola outbreak (Fu, 2002); these results also showed a large increase of infection at the beginning, but once a peak was hit, there was a staggering decrease in infection rates. With this being our basis for how disease should spread throughout the population, we will now apply these findings to our own model.

Methods and Assumptions

In our model, we will have four main groups: vaccinated and wearing a mask, vaccinated and not wearing a mask, unvaccinated and wearing a mask, unvaccinated and not wearing a mask. The goal of having these four groups is to examine all possible combinations of the preventative measures: which are being vaccinated and wearing a mask. It is important to note that we assume a person in the vaccination category to be fully vaccinated. We also assume that a

person who is wearing a mask is wearing some form of a surgical or cloth mask, as it does have a different efficacy rate than a N95 mask (Rettner, 2021). We will calculate the rate of how likely a person is to be infected depending on what preventative measures they have put in place for themselves. We will also observe how infection travels throughout a classroom setting and see who, if anyone, is able to avoid becoming infected. This will allow us to gain a better understanding of the efficacy rates of these different preventative measures. It will also allow us to see how the decisions of one person can impact the health of others.

As of August 19th, 2021, the total population of the United States was (approximately) 332,649,864. Out of this total population, about 46.88% of the population was fully vaccinated. Using the data collected by the CDC, we can see trends that have occurred throughout the COVID-19 pandemic broken down by age category. The age group most likely to test positive for COVID-19 are the 18-29 year olds. However, the age group most likely to die from COVID-19 are ages 85+. We can assume that, while the positive cases for 85+ is one of the lowest total percentages, this is the age group that is most likely to have weakened immune systems, making them more susceptible to dying from this disease (CDC, 2021).

Each variable has its own prevention rate in regards to COVID-19, specifically the delta variant. If a person is fully vaccinated, they are about 36% likely to become infected if they are exposed to the disease (CDC, 2021). Alongside this, if a person is to consistently wear a mask around others, they also only have approximately a 25% chance of becoming infected after exposure (Rettner, 2021). Hence, based on various combinations, if a person is fully vaccinated and religious about wearing their mask, the chance of them being infected is limited to about 9%. However, being vaccinated without a mask or wearing a mask without vaccination significantly

increases the likelihood of becoming infected in comparison to someone who takes both of these precautionary measures.

In order to conduct this research, we will make assumptions about the likelihood of the disease spreading. When discussing those who are vaccinated versus unvaccinated, it is assumed that those who are fully vaccinated, and that those who are unvaccinated have had no dosage. When referring to masks, it is assumed that the person is wearing the mask correctly (covering their nose and mouth), wears it in all indoor settings, and at minimum is wearing a cloth/surgical mask. It is also important to note that, for this research design, we are assuming that all groups involved are completely healthy and have no underlying conditions that may impact their transmission rates. For easier readability, we will assign each person a type that represents their preventative disease measures:

Type 1: Masked and vaccinated

Type 2: Unmasked and vaccinated

Type 3: Masked and unvaccinated

Type 4: Unmasked and unvaccinated

Each type is associated with a different rate of infection. The rates of infection are based on data from the CDC, in which they describe the efficacy rates for both masks and vaccinations against the delta strain of COVID-19. From this data, we see that a person has about a 36% chance of infection if fully vaccinated. We also see about a 25% chance of infection if a person is properly wearing a mask. For our model, we assume that the data gathered from the CDC to be true and assign each type of person a fixed rate of infection, regardless of other circumstances:

Type 1: Masked and vaccinated: 9% probability of infection

Type 2: Unmasked and vaccinated: 36% probability of infection

Type 3: Masked and unvaccinated: 25% probability of infection

Type 4: Unmasked and unvaccinated: 80% probability of infection

After assigning probability rates to each group, we also assigned transmission rates while a person is sick. If a person is infected and wearing a mask, they have a 60% probability of transmitting the disease. If a person is infected and not wearing a mask, they have a 100% probability of transmitting the disease. These percentages are based on the data we have collected and will not vary outside of the guidelines we have set (mask vs no mask). We have also set parameters for how long a person is infectious, which is 3 days, and how long it takes a person to shed the virus, which again is 3 days.

For this project, we will use a classroom setting to model the transmission of COVID-19. In this classroom setting, we have 5 rows and 6 columns, totalling to 30 subjects. On any given day, a subject is sitting in close proximity to a minimum of 3 people and a maximum of 8 people. Our model randomly assigns a type to each subject and then calculates the probability of COVID-19 spreading throughout the entire classroom. The program for the model will run 1,000 times to best see what the likelihood is of being infected from COVID-19 in a classroom setting. This model runs for a total of 45 simulated days, which is roughly the same length of time as a nine week period in a classroom setting. The overall model will be run in the program MATLAB.

Based on the different code written for our program in MATLAB, we will be looking at three different results. First, we have code that tells us the average number of people to become infected based on the amount of people in our 5x6 matrix who are both unmasked and unvaccinated; this is our type 4 preventative measure which has an 80% probability of infection. Each time the simulation runs through a 45 day period which is considered one trial. In order to conduct a meaningful average, we ran the simulation for 100 trials. For each possible number of

type 4 people there could be in the matrix (starting with 0 and ending with 30) we will run 100 trials in order to get the average number of people who would become infected based on how many type 4 people we have in our matrix. Each trial will start with one masked, infected person in the (3,3) coordinate. Our results will then be plotted on a graph for further analysis.

We have another piece of code written for our program that looks at how many type 1 people (our masked and vaccinated group) will become infected based on different probabilities of infection. This section of our results will also have 100 trials for each probability with each individual trial consisting of 45 day periods. It will also be starting with 1 individual who is both masked and infected located in the (3,3) coordinate for each trial. For this section, we will look at the probability of infection ranging from 0% to 100%, then see how many people in our matrix become infected. These results will also be plotted on a graph for further analysis.

Finally, we have another section of code for this program that creates a heat map. The heat map will act as a visual aid when we run the program. This aid shows circles in each spot of our matrix, representing the different types of people that we have previously described. For this section of our program, we will see how the infection rate spreads throughout the classroom based on both the type of person and the proximity to others who are infected, which constantly changes. This heat map runs through the 45 day period, representing 1 trial, with each change in color having a specific meaning. For easier understanding, our heat map is color coordinated with the following meanings:

Red: infected and unmasked

Pink: infected and masked

Blue: previously sick, though no longer infectious

Green: vaccinated and masked (Type 1)

Cyan: vaccinated and unmasked (Type 2)

Yellow: unvaccinated and masked (Type 3)

White: unvaccinated and unmasked (Type 4)

There will always be one masked, infected person in the (3,3) coordinate. However, every other spot in the matrix will be randomized. Each color will also be associated with the probability of infection that has been previously defined, which will impact their rate of infection. Our results will be mapped with an image that represents every change that occurs within our classroom matrix over the 45 day trial.

Results

First, we will look at the results for the part of the program that compares the number of people who are unmasked and unvaccinated to the amount of people who get sick. The larger number of people that are unmasked and unvaccinated in our matrix, the larger number of people who become infected over the course of our 45 day period (Figure 1). This shows that when people are unmasked and unvaccinated, not only are they far more likely to become infected, but they also increase the likelihood of infection for those who are taking more precautions. For example, according to our results in Figure 1, when 5 out of 30 people are unmasked and unvaccinated, approximately 15 people become infected with COVID-19 over the course of 45 days. These results not only show how strong the transmission rate of COVID-19 is, but also the importance of taking precautionary measures to limit the possibility of infection for oneself and others.
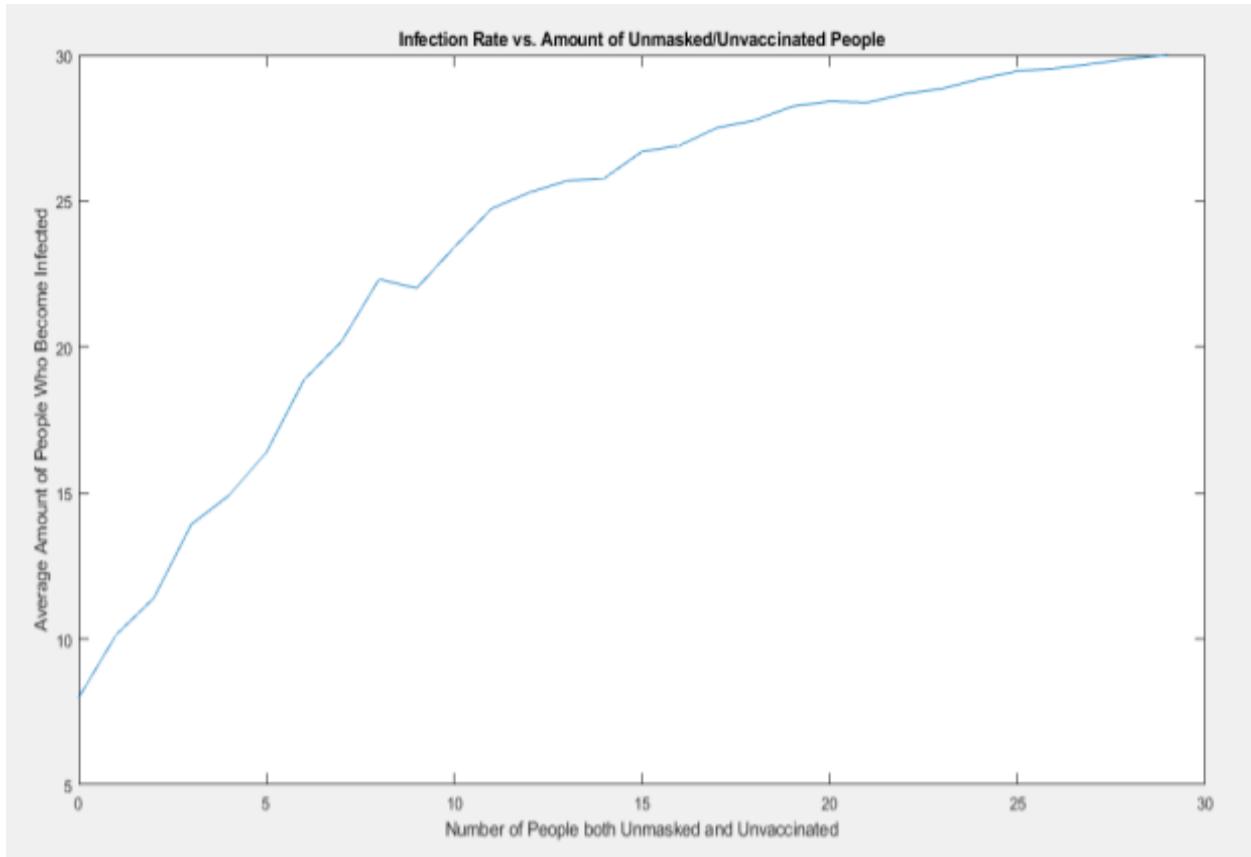
Figure 1. Number of people who become infected during vs. number of people who are both umasked and unvaccinated. As the number of people who are both unmasked and unvaccinated increases within the matrix, so does the number of people who become infected over the course of the simulated 45 day period.

Next, we have our results for the amount of masked, vaccinated people who become infected when we change our rates of infection. We see a strong increase in the number of people who become infected once we start increasing the probability of infection, which is to be expected (Figure 2). However, it is important to note that Figure 2 shows that almost everyone in the matrix will become infected once the probability is increased to the 40-50% range. For these results, it is also important to note that everyone within this matrix is both masked and vaccinated. This tells us that, regardless of how many precautionary measures are taken to prevent disease spread, if there is a high enough probability rate for the disease itself, it is incredibly hard to avoid becoming infected.
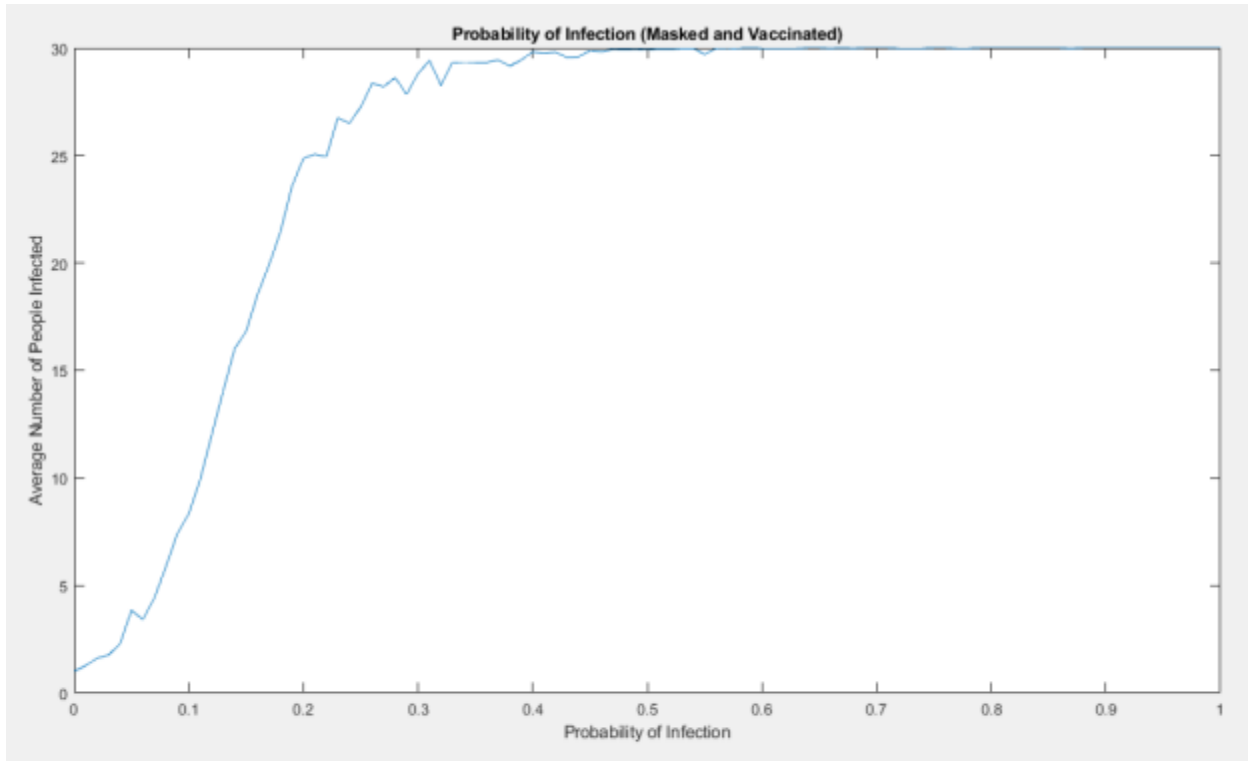
Figure 2. Probability of infection for those who are both masked and vaccinated. As the probability of infection increases, the amount of people who become infected in the matrix increases to the point where, after about a 40% probability of infection, everyone within the matrix will become infected.

Finally, we now will look at our results for the likelihood of becoming infected within our matrix. For this heat map, each trial begins completely randomized except for one position, which always represents a masked, infected person. As we can see in Figure 3, throughout the 45 day trial the status of infection changes pretty rapidly, with the majority of the people becoming infected at some point. Once the trial is over, we see that only 2 out of 30 people have not become infected. There are a few important things to note about the 2 people who did not become infected in trial 1. First, both of these people are masked and vaccinated. They had the lowest probability of infection to begin with due to taking more precautionary measures than some of the others in the matrix. Secondly, the 2 that did not become infected are in ideal positions within the matrix. Both are in the first column, which means that they are not surrounded by others on all sides. Each of the 2 people who did not become infected were only surrounded by at most 4 infected people throughout the trial. In comparison, some people within our matrix become surrounded by 8 infected people throughout the trial. This is similarly reflected in our other trials, which are represented in Figures 4-6. In all of these trials, not only are the ones who did not become infected have the highest level of protection, they were also in spots within the matrix that were not surrounded on all sides by other people. These results not only show how the precautionary measures against COVID-19 play a major role in the likelihood of becoming infected, but also how the location within our matrix plays a role.

Trial 1

Day 1
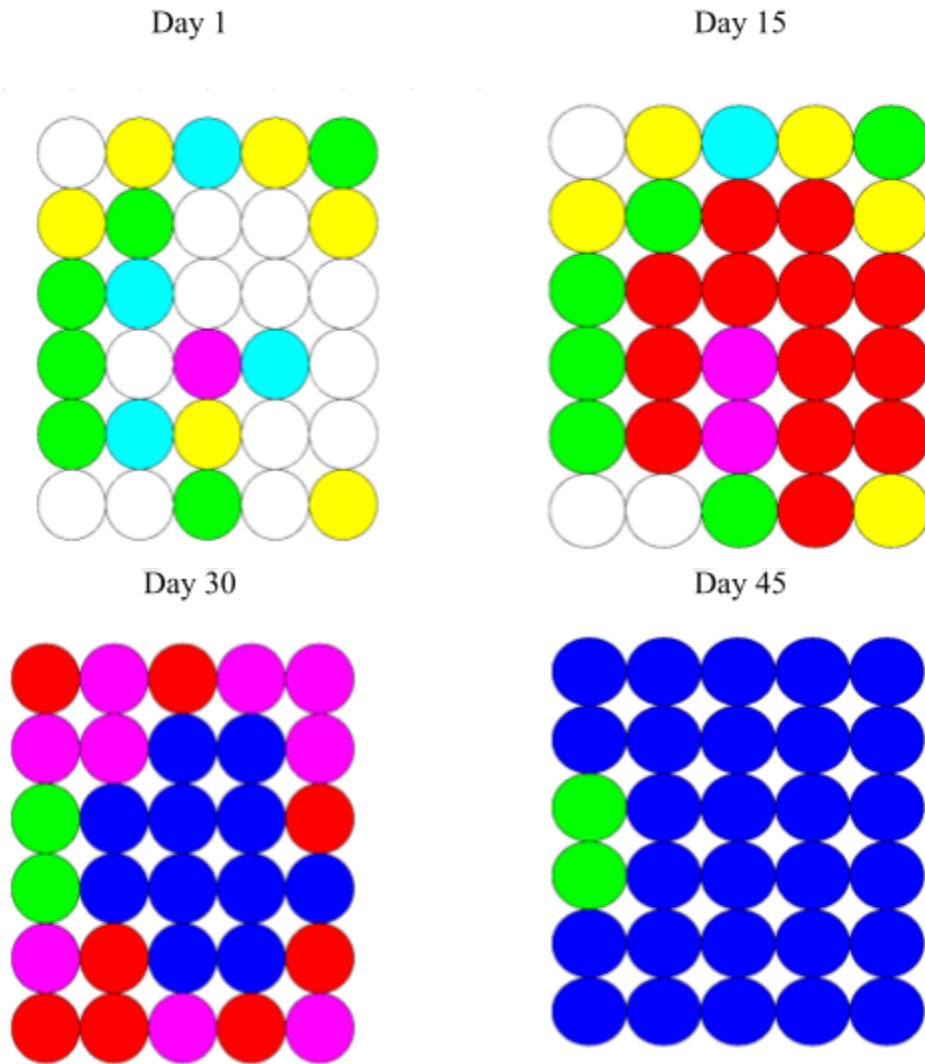
Day 15

Day 30

Day 45

Figure 3. A heat map showing the spread of COVID-19 through a classroom over the course of 45 days in trial 1. On day 1, a single infected and masked individual is seated at position 3,3. Colors represent the initial type of each person and their infection status on day 1, day 15, day 30, and day 45.

Trial 2

Day 1   Day 45

Figure 4. A heat map showing the spread of COVID-19 through a classroom over the course of 45 days in trial 1. On day 1, a single infected and masked individual is seated at position 3,3. Colors represent the initial type of each person and their infection status on day 1, day 15, day 30, and day 45.

Trial 3

Day 1   Day 45

Figure 5. A heat map showing the spread of COVID-19 through a classroom over the course of 45 days in trial 1. On day 1, a single infected and masked individual is seated at position 3,3. Colors represent the initial type of each person and their infection status on day 1, day 15, day 30, and day 45.
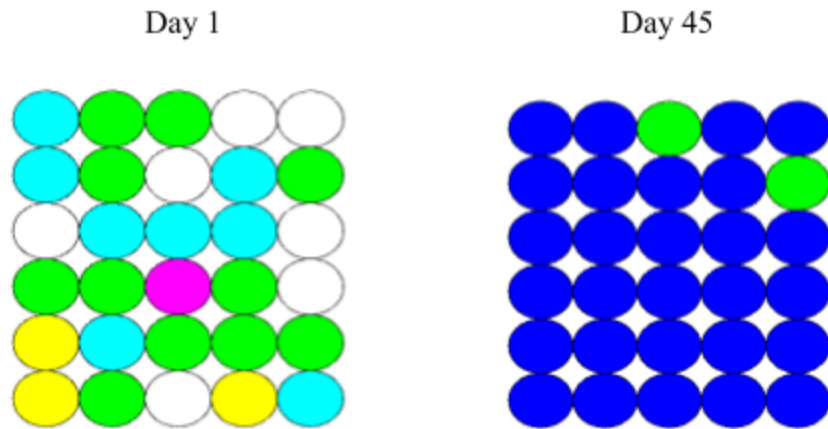
Trial 4

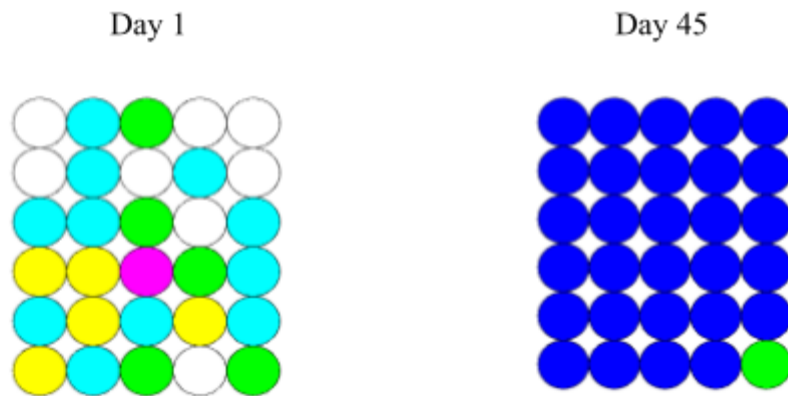Day 1   Day 45

Figure 3. A heat map showing the spread of COVID-19 through a classroom over the course of 45 days in trial 1. On day 1, a single infected and masked individual is seated at position 3,3. Colors represent the initial type of each person and their infection status on day 1, day 15, day 30, and day 45.

Discussion

The results of our cellular automata model show that the more preventative measures one puts in place for themself, the less likely they are to contract COVID-19. However, it is important to note the limitations of our study. One of the most important limitations is that we do not account for any sort of movement within our matrix, which is not realistic. Even in a classroom setting, people tend to be moving around, leaving to go get a drink of water, etc. It also does not consider what happens when a person leaves the classroom and goes on about their day. It would be interesting to use this research as a basis to see how our results might change if those that were in the matrix were moving around within the boundaries, which is much more representative of a real world scenario.

We also do not consider any sort of high risk health conditions or higher susceptibility rates within our model, which again is not very realistic. If everyone in the world was perfectly healthy, this model would be completely accurate. However, that is not the case, and different diseases and immunocompromised statuses can greatly impact one's chances of contracting COVID-19. If we were to continue this research, it would be important to consider how differently the spread of COVID-19 could impact those who already have higher chances of becoming infected because of preexisting conditions. This would allow us to gain a deeper understanding of not only how COVID-19 spreads, but how impactful one person's decision on precautionary measures can impact those around them when there are other underlying conditions at play.

Regardless of these limitations, our results show the importance of using precautionary measures to curb the spread of COVID-19. We also see that our surroundings and the decisions

others make also impact our ability to become infected. While we looked at this on a very small scale, this basic model of cellular automata can not only apply to COVID-19 but to any disease for which the transmission rate is known. Having mathematical models like cellular automata allows us to not only gain further insight into how the disease is spreading, but also how well the protective measures work against the disease. Cellular automata allows us to gain a deeper understanding of COVID-19. By having that deeper understanding, it allows us to be more prepared and better equipped to combat the first pandemic that has been seen in decades.

References

Benneyan, J., Gehrke, C., Ilies, I., & Nehls, N. (2021). Community and Campus

COVID-19 Risk Uncertainty Under University Reopening Scenarios: Model-Based

Analysis. *JMIR Public Health and Surveillance*, *7*(4), e24292.

https://doi.org/10.2196/24292

CDC. (2020, March 28). *COVID Data Tracker*. Centers for Disease Control and

Prevention. https://covid.cdc.gov/covid-data-tracker

Fu, S. C. (n.d.). (2002). *Modelling Epidemic Spread using Cellular Automata*.

Ghosh, S., & Bhattacharya, S. (2021). Computational Model on COVID-19 Pandemic

Using Probabilistic Cellular Automata. *SN Computer Science*, *2*(3), 230.

https://doi.org/10.1007/s42979-021-00619-3

Ramos, A. B. M., & Schimit, P. H. T. (2019). Disease spreading on populations structured

by groups. *Applied Mathematics and Computation*, *353*, 265–273.

https://doi.org/10.1016/j.amc.2019.01.055

Rettner, Rachael. *Face masks may reduce COVID-19 spread by 85%, WHO-backed study

suggests | Live Science*. (n.d.). Retrieved October 2, 2021, from

https://www.livescience.com/face-masks-eye-protection-covid-19-prevention.html

Singhal, T. (2020). A Review of Coronavirus Disease-2019 (COVID-19). *The Indian

Journal of Pediatrics*, *87*(4), 281–286. https://doi.org/10.1007/s12098-020-03263-6

Tupper, P., & Colijn, C. (2021). COVID-19 in schools: Mitigating classroom clusters in the

context of variable transmission. *PLoS Computational Biology*, *17*(7),

e1009120–e1009120. https://doi.org/10.1371/journal.pcbi.1009120

White, S. H., Rey, A. M. del, & Sánchez, G. R. (2007). Modeling epidemics using cellular

  automata. *Applied Mathematics and Computation*, *186*(1), 193–202.

  https://doi.org/10.1016/j.amc.2006.06.126

Wolfram, Stephen. *How Do Simple Programs Behave?: A New Kind of Science | Online by*

  Stephen Wolfram [Page 26]. (n.d.). Retrieved November 14, 2021, from

  https://www.wolframscience.com/nks/p2--an-outline-of-basic-ideas/